Mathematical Biosciences and Engineering

MBE, 19(2): 2095–2119.
DOI: 10.3934/mbe.2022098
Received: 09 October 2021
Accepted: 25 November 2021
Published: 28 December 2021

http://www.aimspress.com/journal/MBE

Research article

Exploring the role of eRNA in regulating gene expression

Heli Tan[1], Tuoqi Liu[2] and Tianshou Zhou[2,3,*]

[1] School of Financial Mathematics and Statistics, Guangdong University of Finance, Guangzhou 510521, China
[2] School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, China
[3] Guangdong Province Key Laboratory of Computational Science, Guangzhou 510275, China

* **Correspondence:** Email: mcsthzsh@mail.sysu.edu.cn.

Abstract: eRNAs as the products of enhancers can regulate gene expression via various possible ways, but which regulation way is more reasonable is debatable in biology, and in particular, how eRNAs impact gene expression remains unclear. Here we introduce a mechanistic model of gene expression to address these issues. This model considers three possible regulation ways of eRNA: Type-I by which eRNA regulates transcriptional activity by facilitating the formation of enhancer-promoter (E-P) loop, Type-II by which eRNA directly promotes the mRNA production rate, and mixed regulation (i.e., the combination of Type-I and Type-II). We show that with the increase of the E-P loop length, mRNA distribution can transition from unimodality to bimodality or vice versa in all the three regulation cases. However, in contrast to the other two regulations, Type-II regulation can lead to the highest mean mRNA level and the lowest mRNA noise, independent of the E-P loop length. These results would not only reveal the essential mechanism of how eRNA regulates gene expression, but also imply a new mechanism for phenotypic switching, namely the E-P loop can induce phenotypic switching.

Keywords: eRNA; enhancer; gene expression model; E-P looping; chemical master equation

1. Introduction

Enhancers, located upstream or downstream from the transcription start sites (TSSs), are important cis-acting elements, which can activate or promote the transcription levels of their target genes, independent of their position and orientation to the target genes [1–5]. It has been reported that there are more than 400,000 enhancers in human genome [6,7], and this number is far more than that of protein-encoding genes. Active enhancers can recruit a number of transcription factors (TFs) and

RNAPII to control precise gene expression in tissue-specific patterns required for the directional development of cell types and lineages [8–12].

Biological experiments have reported fundamental characteristics of enhancers: (1) enhancers with the promoters of protein-encoding genes can form chromatin loops, thus producing a large amount of eRNAs [13,14]; (2) Functionally active enhancers can be transcribed into different types of eRNAs in an unidirectional or a bidirectional manner [9,15–19]; (3) eRNAs produced by enhancers are extremely unstable because they are quickly degraded by the RNA exosome, a ribonucleic acid exonuclease complex involved in the degradation and processing of intracellular RNAs [20–23]; (4) The eRNA abundance is so low that eRNAs are often viewed as the byproducts of enhancer transcription [24]. These experimental facts provide guidelines for modeling and analysis of gene expression regulated by eRNAs, but biological functions of the eRNAs remain elusive.

Two theories have been proposed to interpret the roles of eRNAs. The first theory is the so-called "transcription noise theory", which considers that eRNAs are merely the byproducts of enhancer transcription without functional significance. Some biologists supported this idea, e.g., Albin Sandelin thought that enhancer and the regulatory promoter may form a stable DNA loop when high concentrations of RNAP and TFs accumulate [25]. The excess RNAPII machinery collides stochastically with enhancers to initiate transcription [18–26]. The second theory is the so-called "functional theory", which considers that eRNAs themselves may play a role of activating gene expression since transcription does not seem a random phenomenon but would be controlled by various factors [27–33]. For example, stationary enhancers have no transcriptional activities, meaning that chromatin accessibility is decided by DNase I hypersensitivity [34]. With regard to the "functional theory", several hypotheses have been made to describe the role of eRNAs at distinct stages: (1) Enhancer-promoter (E-P) looping. For example, in human breast cancer cells, the eRNAs expressed from estrogen receptors that bind enhancers promote specific E-P interactions by recruiting adhesion proteins to the enhancers [27]. Also for example, in prostate cancer cells, the eRNAs expressed from the Kallikrein-related peptidase 3 enhancer accelerate the interaction between the KLK3 enhancer and the KLK2 promoter, by forming a complex with the androgen receptor and subunit of Mediator complex, MED1 [29]; (2) Chromatin remodeling. During the myogenic differentiation of C2C12 skeleton muscle cells, eRNAs expressed from MYOD1 core enhancer facilitate the accessibility of chromatin and enhance the occupancy of RNAPII to the promoter region and the subsequent gene expression [31]; (3) Early transcription elongating. In the induction of immediate early genes in neuronal cells, eRNAs facilitate the efficient release of negative elongation factors (NELF) from their target gene promoters through direct association with the RNA recognition motif in the NELF-E subunit, thus promoting the synchronous expressions of the target genes [33]. A schematic diagram for the interaction between enhancer and promoter is shown in Figure 1. Which of the above three hypotheses is more reasonable is debating. In particular, the mechanism of how eRNAs affect gene expression is unclear.

So far, many molecular mechanisms have been proposed to explain the impact of eRNAs on gene expression, but to our knowledge, there seem no theoretical models to investigate eRNAs' potential functions. In this paper, we build a mechanistic model of stochastic gene expression, which considers both an E-P loop characterizing the E-P interaction and different feedbacks representing different ways that eRNAs impact gene expression. Although simple, this model still captures fundamental events taking place in gene expression, such as random synthesis and degradation of eRNA and mRNA, and stochastic switching between promoter activity states as well as eRNA regulation of the mRNA

expression. It also considers three possible regulation ways: Type-I regulation by which eRNAs regulates transcriptional activity by facilitating the formation of the E-P loop's formation, Type-II regulation by which eRNAs directly increase the production rate of mRNA, and the combination of these two ways (called mixed regulation). Note that the "transcription noise theory" does not consider regulation.

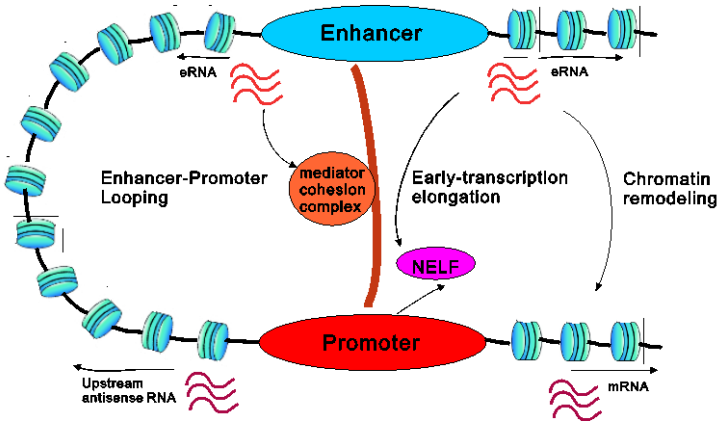

Figure 1. Schematic of enhancer-promoter interactions, which are reflected in the following three stages. (1) E-P looping: eRNAs can associate with the subunits of mediator-cohesion complex; (2) Chromatin remodeling: eRNAs can remodel chromatin structure and increase the accessibility of RNAPII machinery, thus improving transcription; (3) Early transcription elongation. Some eRNAs can facilitate transient release of the negative elongation factors, helping the RNAPII enter into the TSSs and initiate a productive elongation stage.

We analytically derive approximate steady-state mRNA and eRNA probability distributions as well as their statistical quantities such as mean and expression noise (defined as the ratio of the variance over the squared mean). Interestingly, we find that with the increase of the E-P loop length, the mRNA probability distribution can switch from unimodality to bimodality and then be back to unimodality in all the three cases of eRNA regulation. We also find that in contrast to the other two regulations, the Type-II regulation can result in the highest mean mRNA level and the lowest mRNA noise, independent of the E-P loop length. Since these results are qualitative, we conclude that the E-P loop can induce phenotypic switching, which may be utilized by organisms or live cells. Our analysis validates the eRNA's "functional theory", whereas our results would have potential biological implications.

2. Model description

2.1. Mathematical model

Although the interaction between the enhancer and the promoter of a gene would be diverse (in particular, in eukaryotic cells), here we assume that they form a loop (a representative chromatin structure) so that RNAPII can access a specific binding site to initiate transcription. This assumption

is reasonable. In fact, experimental evidences indicated that the formation of a DNA loop is beneficial to structural stability and thus to recruitment of TFs and RNA polymerases. For example, E. Savitskaya et al. verified by experiment that long-distance functional interactions between Su (Hw) insulators that can regulate the E-P communication in Drosophila melanogaster [35]. Looping is critical for process ranging from gene regulation to recombination and repair, and the boundaries of loop domains (insulators) are the determinants of chromosome form and function in multicellular eukaryotes [36]. M. Julien et al. suggested that the genome segment delimited by topological insulators and comprising the gene and its enhancer regulatory sequences is the plausible adaptor, which mediating the correspondence between the concentrations of specific transcription factors in the cell nucleus and gene expression level [37]. In addition, P. Charles et al. proposed that the Activity-by contact model of enhancer-promoter regulation from thousands of CRISPR perturbations to map and predict which enhancers regulate which genes [38]. We also assume that eRNA and mRNA are expressed only when the E-P loop is formed. This assumption is also reasonable since a promoter must recruit TFs and RNA polymerases to initiate transcription. Using the idea of the mapping method proposed in ref. [39], we build a mechanistic model of stochastic gene expression to model the effect of eRNA on gene expression, referring to Figure 2. This model is, in essence, similar to the extensively studied ON-OFF model of stochastic transcription [40–44]. That is, the gene promoter has one active (ON) state that corresponds to the case that the E-P loop is formed so that RNAPII can access a specific binding site to initiate transcription, where DNAs are transcribed into mRNAs and enhancers are transcribed into eRNAs, and one inactive (OFF) state that corresponds to the case that the E-P loop is not formed, where transcription is not existent, namely, neither the gene nor the enhancer is expressed. Furthermore, stochastic transitions exist between ON and OFF states.

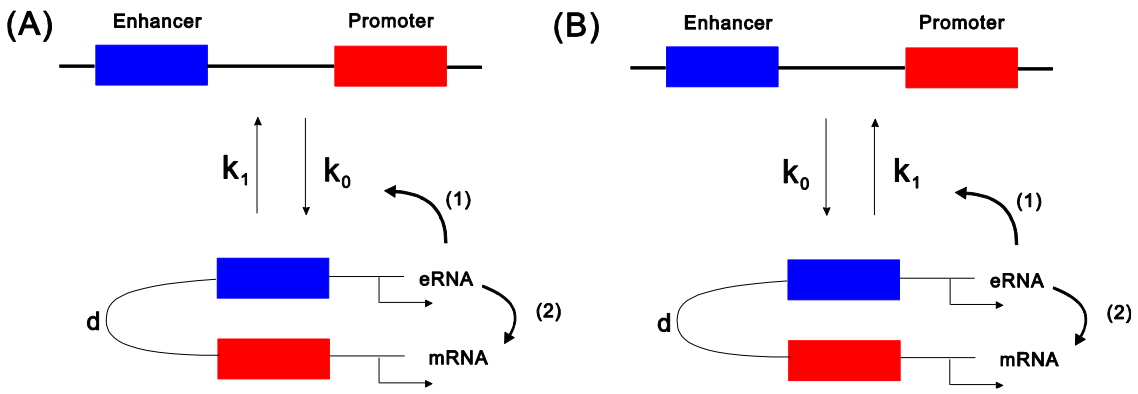

Figure 2. Schematic of gene models considering eRNA and its regulation as well as its interaction with the promoter. Neither the gene nor the enhancer is expressed if the enhancer and the promoter do not form a DNA loop and both are expressed otherwise. In (A) or (B), eRNAs produced by the enhancer are assumed to regulate gene expression in three possible manners: Type-I regulation, Type-II regulation and mixed regulation.

In addition, the model also considers feedback. Usually, mRNA and eRNA need to use the same resource, polymerases, which, in principle, may couple two populations of the molecules. If a polymerase pool is sufficiently large, then we can neglect the mRNA-eRNA feedback bridged by polymerase. The produced eRNAs then serve as transcriptional factors to regulate the target gene

expressions. Based on experimental evidences, we introduce three possible regulation ways (or mechanisms) of eRNA: Type-I regulation, Type-II regulation and mixed regulation, as described in the introduction. Note that the first regulation mechanism corresponds to the first hypothesis in the "functional theory", whereas the second regulation mechanism integrates the latter two hypotheses in the "functional theory". Nevertheless, it is unclear which of the three regulation mechanisms is more possibly utilized by live organisms or cells.

For clarity, we list all biochemical reactions corresponding to Figure 2A as follows

$$\begin{aligned} &I \xrightarrow{k_0+\alpha[eRNA]} A,\ A \xrightarrow{k_1} I,\ A \xrightarrow{\mu_1+\beta[eRNA]} A+mRNA, \\ &A \xrightarrow{\mu_2} A+eRNA,\ \ mRNA \xrightarrow{\delta_1} \varnothing, eRNA \xrightarrow{\delta_2} \varnothing, \end{aligned} \tag{1}$$

where k_0 and k_1 are respectively transition rates from inactive (I) to active (A) states and vice versa in the absence of eRNA regulation, μ_1 and μ_2 are respectively transcription rates, δ_1 and δ_2 are respectively degradation rates of mRNA and eRNA. In Eq 1, $[eRNA]$ represents the number of eRNA molecules or the eRNA concentration, α and β represent two feedback strengths and are assumed as nonnegative constants. If $\alpha>0$ but $\beta=0$, then the model reduces to the case of DNA looping. If $\alpha=0$ but $\beta>0$, then it reduces to the case of Type-II regulation. If $\alpha>0$ and $\beta>0$, then it corresponds to the case of mixed regulation. If $\alpha=0$ and $\beta=0$, then it corresponds to the case of no regulation. Thus, we can study the effect of eRNA on the mRNA expression in a united framework.

Let m and n represent the numbers of mRNA and eRNA molecules, respectively. To better trace the time evolution of probability distributions of mRNA and eRNA molecule numbers, we introduce two distribution functions $P_I(m,n;t)$ and $P_A(m,n;t)$, which represent the probabilities that mRNA have m molecules and eRNA have n molecules at time t when the gene is respectively at I and A states. Thus, the total probability, denoted by $P(m,n;t)$, is given by $P=P_A+P_I$. Then, the chemical master equation corresponding to Eq 1 takes the following form

$$\begin{aligned} \frac{\partial P_I(m,n;t)}{\partial t} &= k_1 P_A(m,n;t) - (k_0+\alpha n) P_I(m,n;t) + \delta_1 (E_m - S)\left[mP_I(m,n;t)\right] \\ &\quad + \delta_2 (E_n - S)\left[nP_I(m,n;t)\right], \\ \frac{\partial P_A(m,n;t)}{\partial t} &= -k_1 P_A(m,n;t) + (k_0+\alpha n) P_I(m,n;t) + \delta_1 (E_m - S)\left[mP_A(m,n;t)\right] \\ &\quad + \delta_2 (E_n - S)\left[nP_A(m,n;t)\right] + (\mu_1+\beta n)\left(E_m^{-1} - S\right)\left[P_A(m,n;t)\right] \\ &\quad + \mu_2 \left(E_n^{-1} - S\right)\left[P_A(m,n;t)\right], \end{aligned} \tag{2}$$

where $m,n=0,1,2,\cdots$, S is the unit operator, and E_l is the common step operator and E_l^{-1} is its

inverse (where $l=m$ or n). The initial condition is set as $P_I(0,0;0)=1$ and $P_A(0,0;0)=0$. We are interested in stationary eRNA and mRNA distributions. Note that at steady state, Eq 2 becomes

$$\begin{aligned}
&k_1P_A(m,n)-(k_0+\alpha n)P_I(m,n)+\delta_1(E_m-S)\left[mP_I(m,n)\right]+\delta_2(E_n-S)\left[nP_I(m,n)\right]=0,\\
&-k_1P_A(m,n)+(k_0+\alpha n)P_I(m,n)+\delta_1(E_m-S)\left[mP_A(m,n)\right]+\delta_2(E_n-S)\left[nP_A(m,n)\right]\\
&\quad+(\mu_1+\beta n)\left(E_m^{-1}-S\right)\left[P_A(m,n)\right]+\mu_2\left(E_n^{-1}-S\right)\left[P_A(m,n)\right]=0.
\end{aligned} \tag{3}$$

Introduce two factorial marginal distributions for eRNA and mRNA, which are respectively defined as $Q_i(n)=\sum_m P_i(m,n)$ and $R_i(m)=\sum_n P_i(m,n)$, where $i=I,A$. Thus, the total stationary distributions of eRNA and mRNA are $Q(n)=Q_I(n)+Q_A(n)$ and $R(m)=R_I(m)+R_A(m)$, respectively. One main aim of this paper is to derive the analytical expressions of $Q(n)$ and $R(m)$.

2.2. Dependence of looping rate on E-P loop length

Apart from the eRNA's regulation mechanism, the E-P looping rate is also an important factor affecting gene expression. Since an enhancer's position relative to its promoter is not fixed, the E-P looping rate is assumed to change in a finite range. Note that the transition rate from OFF to ON states in Figure 2, k_0, actually represents the rate that the enhancer and the promoter form a DNA loop (denoted by k_{loop}), i.e., $k_0=k_{loop}$. In order to quantify how the E-P loop length (denoted by d) impacts gene expression, it is needed to know the dependence of k_0 on d. For a single DNA loop, previous works studied the impact of the E-P loop length on the looping rate, and even gave empirical formulae [47,48]. In our case, the corresponding formula takes the form of $k_{loop}=k_{on}\exp\left(-\frac{u}{d}-v\log d+wd+z\right)$, which is motivated by the worm-like chain model of DNA bending [49], where k_{loop} represents the rate of the DNA looping, and d represents the E-P loop length along the DNA line, i.e., the distance between the promoter and the enhancer. Accordingly, some parameter values can be set as $k_{on}=1,u=140.6,v=2.52,w=0.0014,z=19.9$, which are obtained by fitting experimental data [47,48]. As mentioned on the Introduction, enhancers located upstream or downstream of the target gene. For target gene, the different position of enhancer will directly affect the E-P looping rate k_{loop}. In other words, once the target gene and the enhancer are fixed, then the k_{loop} is determined. However, for different target gene, the position of enhancer is different, so k_{loop} is also different. Under this setting, given the value of k_0, we can determine the

DNA loop length (also called distance) d from the above empirical formula. However, the case of $k_1 = k_{loop}$ is also possible in theory. Therefore, based on the above empirical formula, we can also determine d if the value of k_1 is given.

As is well known, enhancers' regulation has a long-distance effect. Although enhancers are cis-acting elements, the position relative to their target genes may be highly changeable. Generally, an enhancer is located ~$200\,bp$ upstream from the target promoter, but can also facilitate the transcription activity of a distant promoter even with the distance being >$10\,kb$ [3,50]. In the following sections, we will set the E-P loop length between $200\,bp$ and $8000\,bp$ in our numerical simulations, to obtain qualitative results.

Finally, in this section, we point out: (1) Most eRNA's have been proposed to have very short half-lives (~7.5 mins) and only a few have long half-lives [33], implying that the degradation rate of eRNA is in general larger than that of mRNA. In our simulation, we will set $\delta_1 = 1$ (one unit) and $\delta_2 = 5$. (2) The amount of eRNAs is usually much less than that of mRNAs [51], implying that the production rate of the former is much smaller than that of the latter. In our simulation, we will set $\mu_1 = 20$ and $\mu_2 = 2$. (3) The choice of other parameter values is based on the setting of parameter values used often in the study of the common ON-OFF model at the transcription level [40].

3. Results

First, there are two possible ways to determine the DNA loop length (d) if the other parameters are fixed: increasing k_0, and reducing k_1. Second, we numerically find that in whichever way (i.e., given the value of k_0 or k_1, which determines d according to the above empirical formula), the qualitative impact of increasing k_0 or reducing k_1 on the mRNA expression is the same (referring to an example shown in the Appendix of this paper). Therefore, we will only show results in the case that d is determined by the previous setting of k_0 that is allowed to change in a finite range (implying that d changes in another finite range).

3.1. Analytical distributions and statistics

Analytical results can often provide the comprehensive understanding of a system's behavior. Here we present analytical results for eRNA and mRNA distributions, and statistical quantities such as expectations and variances as well as noise intensities. The detailed mathematical derivations are put in Appendix A of this paper.

First, consider the eRNA case. We can show that the exact, analytical eRNA distribution is given by

$$Q(n)=\frac{\tilde{\mu}_2^n}{n!}\frac{\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1}\right)_n}{\left(\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1}\right)_n}\frac{{}_1F_1\left(n+\frac{\tilde{k}_0}{\tilde{\alpha}+1},n+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)}{{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)}, \tag{4}$$

where all the parameters are normalized by δ_2, i.e., $\tilde{\mu}_2=\mu_2/\delta_2$, $\tilde{k}_0=k_0/\delta_2$, $\tilde{k}_1=k_1/\delta_2$ and $\tilde{\alpha}=\alpha/\delta_2$, function ${}_1F_1(a,b;z)$ is a confluent hypergeometric function [52], defined as ${}_1F_1(a,b;z)=\sum_{n=0}^{\infty}\frac{(a)_n}{(b)_n}\frac{z^n}{n!}$, and $(c)_n$ is the Pochhammer symbol, defined as $(c)_n=\Gamma(c+n)/\Gamma(n)$ with $\Gamma(\cdot)$ being the common Gamma function. Equation 4 indicates that the stationary eRNA distribution is independent of mRNA. Thus, the eRNA noise intensity is given by

$$\eta_n^2=\frac{\langle n^2\rangle-\langle n\rangle^2}{\langle n\rangle^2} \tag{5}$$

where

$$\langle n\rangle=\frac{\tilde{\mu}_2\tilde{k}_0}{\tilde{k}_0+\tilde{k}_1}\frac{{}_1F_1\left(1+\frac{\tilde{k}_0}{\tilde{\alpha}+1},1+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\tilde{\mu}_2\left(1+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}\right)\right)}{{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)}, \tag{5a}$$

$$\langle n^2\rangle=\frac{\tilde{\mu}_2\tilde{k}_0}{2\left(\tilde{k}_0+\tilde{k}_1\right)}\frac{{}_1F_1\left(2+\frac{\tilde{k}_0}{\tilde{\alpha}+1},2+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\tilde{\mu}_2\left(1+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}\right)\right)}{{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)}. \tag{5b}$$

Then, consider the stationary mRNA distribution. Under some assumed conditions, we can derive the approximate expression of this distribution (seeing Appendix A for derivation)

$$R(m)=\frac{\left(\tilde{\mu}_1+\tilde{\beta}\langle n\rangle\right)^m}{m!}\frac{\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle\right)_m}{\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle+\tilde{k}_1\right)_m}{}_1F_1\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle+m,\tilde{k}_0+\tilde{\alpha}\langle n\rangle+\tilde{k}_1+m;-\left(\tilde{\mu}_1+\tilde{\beta}\langle n\rangle\right)\right) \tag{6}$$

where $\langle n\rangle$ is given by Eq 5a. This analytical expression shows how eRNA quantitatively impacts the stationary mRNA distribution. Apparently, this impact is nonlinear. We point out that this approximate distribution has been verified by the Gillespie stochastic simulation algorithm [53], referring to Figure 3A where the E-P loop length is set as $1500bp$. From the panels of this figure, we observe that in all the three cases of eRNA regulation, the absolute differences between theoretical and

numerical results do not exceed 0.01, indicating that our approximate method (seeing Appendix A for detail) is effective.

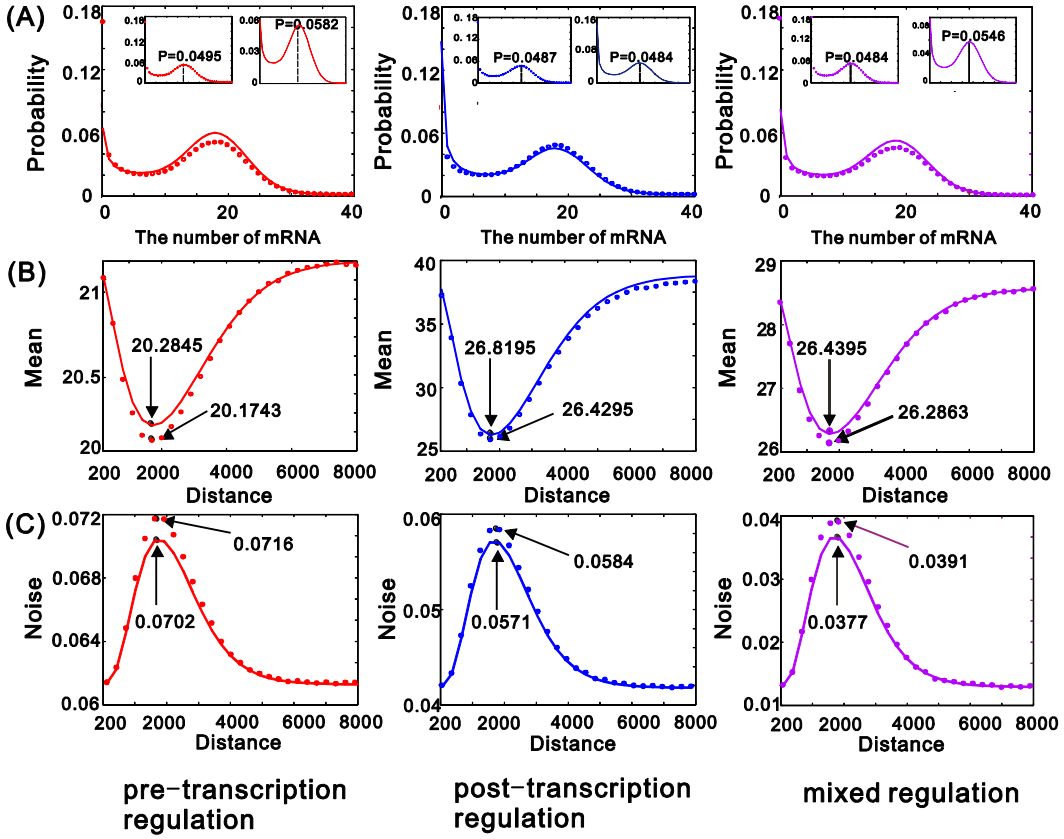

Figure 3. Shown are the results (solid lines) obtained by theoretical prediction and the results (circles) obtained by the Gillespie algorithm in the cases of Type-I regulation (red), Type-II regulation (blue), and mixed regulation (purple). (A) mRNA probability distributions, where the probabilities corresponding to the second peak (i.e., the peak away from the origin) are shown in the small boxes. The E-P loop length is set as $d = 1500bp$. (B) Dependence of the mRNA mean on d, where the smallest means are indicated by arrows. (C) Dependence of the mRNA noise on d, where the largest noise intensities are indicated by arrows. In (A,B,C), parameter values are set as: $k_1 = 0.2, \mu_1 = 20$, $\mu_2 = 2, \delta_1 = 1, \delta_2 = 5$ and feedback strength: $\alpha = 1$ for Type-I regulation, $\beta = 1$ for Type-II regulation, and $\alpha = 0.5, \beta = 0.5$(A) but $\alpha = 0.7, \beta = 0.3$(B,C) for mixed regulation. In (B,C), the E-P loop length changes from 200 to $8000bp$.

The mean mRNA is given by

$$\langle m\rangle = \left(\tilde{\mu}_1 + \tilde{\beta}\langle n\rangle\right)\frac{\tilde{k}_0 + \tilde{\alpha}\langle n\rangle}{\tilde{k}_0 + \tilde{\alpha}\langle n\rangle + \tilde{k}_1}, \tag{7a}$$

whereas the mRNA noise intensity by

$$\eta_m^2 = \frac{\tilde{k}_0 + \tilde{\alpha}\langle n\rangle + \tilde{k}_1}{\left(\tilde{\mu}_1 + \tilde{\beta}\langle n\rangle\right)\left(\tilde{k}_0 + \tilde{\alpha}\langle n\rangle\right)} + \frac{\tilde{k}_1}{\left(\tilde{k}_0 + \tilde{\alpha}\langle n\rangle\right)\left(\tilde{k}_0 + \tilde{\alpha}\langle n\rangle + \tilde{k}_1 + 1\right)}. \tag{7b}$$

Both show how the mean of eRNA molecules affects the mean and noise intensity of mRNA. Moreover, $\langle m\rangle$ is a monotonically increasing function of $\langle n\rangle$ whereas η_m^2 is a monotonically decreasing function of $\langle n\rangle$. Similar to the case of mRNA distribution, the approximate mean and noise intensity of mRNA have also been verified by the Gillespie stochastic simulation algorithm [53], referring to Figure 3B (for the mRNA mean) and 3C (for the mRNA noise intensity). In fact, the maximum error between the two mRNA means does not exceed 0.4, whereas the maximum error between the two mRNA noise intensities does not exceed 0.002. Therefore, except that the error is slightly larger at $d = 2000bp$ (no more than 0.5), the two methods are still in good agreement in other E-P looping length. These results further indicate that our approximate method is reasonable.

3.2. Impact of E-P loop length on mRNA probability distribution

Experimental evidences have indicated that an enhancer may be located in any position of the promoter's upstream or downstream. The enhancer can not only prompt the transcription of the core promoter but can also facilitate that of any promoter near it. Therefore, it is necessary to investigate how the E-P loop length impacts gene expression and further cell-to-cell variability in isogenic cells. Here we address this issue in terms of mRNA distribution, with reasons stated as follows. First, in contrast to a single stochastic trajectory, the distribution can provide more complete information on stochastic dynamics of the underlying system. Second, distributions provide an intuitive understanding for cellular phenotypes and phenotypic diversity as well as phenotypic switching [54–56].

In order to show the effect of the E-P loop length on the mRNA distribution, we consider four cases: Type-I regulation, Type-II regulation, mixed regulation and non-regulation. Numerical results are demonstrated in Figure 4, where the lines represent the results obtained by the Gillespie algorithm [53] whereas the circles represent the results obtained by theoretical prediction. We observe that with the increase of the E-P loop length, the mRNA distribution can change from unimodality to bimodality and then be back to unimodality in all the cases, independent of eRNA regulatory patterns.

In Figure 4A, the first, second, third and fourth columns correspond respectively to Type-I regulation, mixed regulation, Type-II regulation and non-regulation. In all the cases, the E-P loop length can result in unimodal or bimodal mRNA distribution where one peak is close to zero and the other peak is away from zero. Specifically, the mRNA distribution is unimodal for a small d (for example, d is set as $200\,bp$) in all the four cases of eRNA regulation, referring to the first row in Figure 4A. At least d is about $600\,bp$ is needed to generate bimodality for Type-I regulation, at least

d is about $800\,bp$ for mixed regulation, at least d is about $1000\,bp$ for Type-II regulation, and at least d is about $1200\,bp$ for non-regulation, referring to the second row in Figure 4A. With the increase of d, the bimodal mRNA distribution gradually become unimodal, but the E-P loop length needed for this change is different among the four cases. Specifically, at least d is about $3500\,bp$ is needed for Type-I regulation, at least d is about $4600\,bp$ for mixed regulation, at least d is about $5500\,bp$ for Type-II regulation, and at least d is about $7000\,bp$ for non-regulation, referring to the fourth row in Figure 4A. The panels from the second row to the third row in Figure 4A show how the E-P loop length impacts the shape of the bimodal mRNA distribution in all the four cases of eRNA regulation. Figure 4B demonstrates the time series of the mRNA number in the four cases of eRNA regulation: from top to bottom, the fifth row for bimodality and the sixth row for unimodality. In order to help understand the mechanism of how the E-P loop length induces the transition between ON and OFF states in all the four cases of eRNA regulation, we also plot the time evolution of promoter activity, referring to Figure 4C.

In a word, we numerically find that the shortest E-P loop length is needed to generate bimodality or to switch from bimodality to unimodality in the case of Type-I regulation, but the longest d is needed in the case of non-regulation. In the other two regulation cases, the d needed for bimodality or switching is in between the shortest and longest E-P loop lengths. The results shown in Figure 4 would have important biological implication. Perhaps cells would make use of the E-P loop for better survival in complex environments.

Finally in this subsection, we give a simple interpretation for the appearance of bimodality. First, while the common on-off model of gene expression itself has the potential to generate bimodality, our model has the similar one [40,41]. Second, eRNA in our gene model is designed to regulate the expression of the gene in an activating manner. Due to the low production rate and high degradation rate of eRNA, the number of eRNA is generally small, but this activating effect on gene expression indeed exists and is unneglectable, and even can induce phenotypic switching, as demonstrated in Figure 4. In addition, we point out that previous studies showed that a simple genetic auto-activation loop can lead to bistability (meaning that the underlying deterministic system has two stable equilibria) but can also lead to bimodality in the stochastic case [57]. In contrast, we showed that the E-P loop can induce bimodality. Therefore, our mechanism is different from the previous one.

3.3. Effects of the E-P loop length on the mRNA mean

Figure 5 shows how the E-P loop length affects the mean mRNA level. Specifically, Figure 5A demonstrates the dependence of the mRNA mean on the E-P loop length in all the four cases of eRNA regulation, Figure 5B demonstrates the dependence of the ratio of the mRNA mean in all the three cases of eRNA regulation on the E-P loop length, Figure 5C demonstrates the impact of the E-P loop length on transition rate k_0, and Figure 5D demonstrates the dependence of mRNA mean on both E-P loop length and feedback strength.

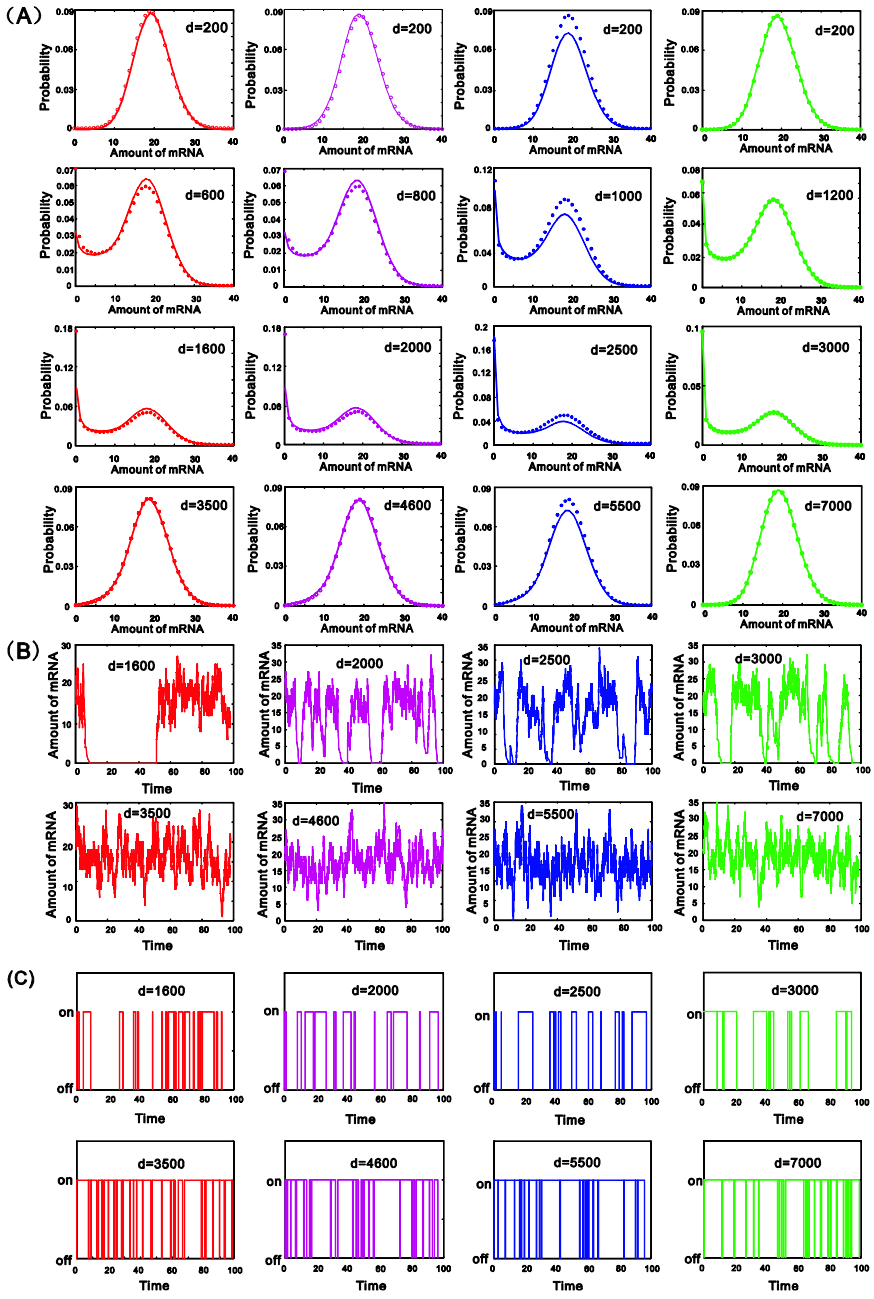

Figure 4. Probability distribution of mRNA and time evolution of the mRNA number as well as the time evolution of the promoter activity, where red lines (first column) correspond to Type-I regulation, purple lines (second column) to mixed regulation, blue lines (third column) to Type-II regulation, and green lines (fourth column) to non-regulation. The solid lines represent the results obtained by theoretical prediction, whereas the circles represent the results obtained using the Gillespie algorithm. (A) Dependence of mRNA probability distribution on the E-P loop length. (B) Time evolution of the mRNA number. (C) Time evolution of the promoter activity. The parameter values are set as: $k_1 = 0.2, \mu_1 = 20, \mu_2 = 2, \delta_1 = 1, \delta_2 = 5,$ and feedback strengths are set as: $\alpha = 0.8$ for Type-I regulation, $\alpha = 0.4, \beta = 0.4$ for mixed regulation, and $\beta = 0.8$ for Type-II regulation.

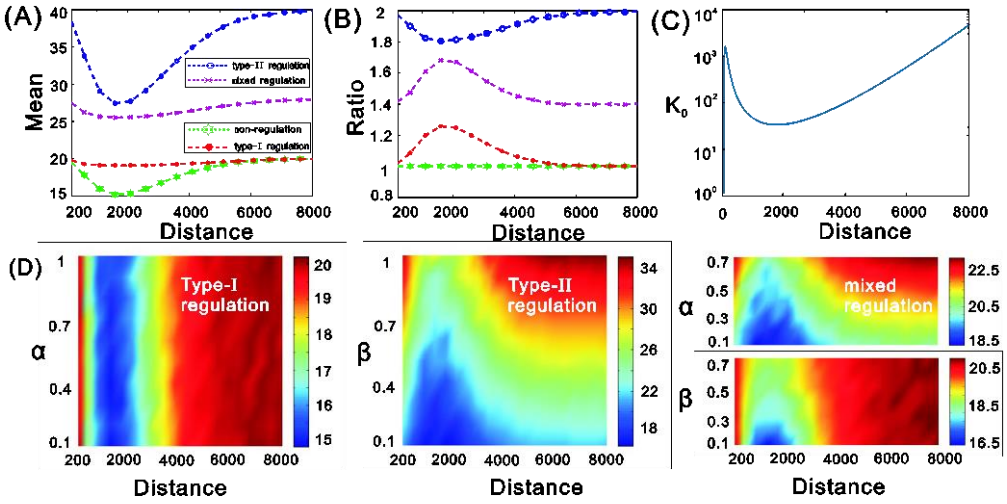

Figure 5. Influence of eRNA on the mRNA mean. (A) Dependence of the mRNA mean on the E-P loop length, where red dotted dash lines correspond to Type-I regulation, purple forked dash lines to mixed regulation, blue circled dash lines to Type-II regulation, and green hexagonal star dash lines to non-regulation. (B) The ratio of the mRNA mean in three cases of eRNA regulation over that in case of non-regulation. (C) Dependence of k_0 on the E-P loop length. (D) Dependence of mRNA mean on both E-P loop length and feedback strength. Parameter values are set as: $k_1 = 1, \mu_1 = 20, \mu_2 = 2, \delta_1 = 1, \delta_2 = 5, \alpha = 0.6,$ $\beta = 0.4$, and the E-P loop length changes from $200\,bp$ to $8000\,bp$.

From Figure 5A, we observe that the curve in the case of non-regulation is always below the curves in three cases of eRNA regulation for all possible E-P loop lengths that are assumed to change in the interval from $200\,bp$ to $8000\,bp$ in our simulation, and the curve in the case of Type-II regulation is above that in the case of mixed regulation, which is above that in the case of Type-I regulation. Moreover, the changing range of the mean mRNA level is larger in the case of Type-II regulation than that in the cases of Type-I and mixed regulations, indicating that the Type-II regulation can amplify the mRNA expression level in contrast to the other two ways of eRNA regulation (i.e., Type-I and mixed regulations). We also observe that the mRNA mean first decreases for smaller E-P loop lengths and then increases with the increase of the E-P loop length, that is, the mean mRNA level reaches the minimum in all the four cases of eRNA regulation.

Figure 5B demonstrates the dependence of the mean mRNA ratio (defined as the ratio of the mean mRNA in the case of eRNA regulation over the mean mRNA in the case of non-regulation) on the E-P loop length for three ways of eRNA regulation, where by "relative" we mean that the mean mRNA level in the case of non-regulation is set as unit (or the mean mRNA level in the case of non-regulation is taken as a reference). We observe that these relative changes exhibit different characteristics, depending on the way of eRNA regulation. Specifically, the mean mRNA ratio reaches the minimum in the case of Type-II regulation but the maximum in both cases of Type-I and mixed regulations. In

addition, for all possible E-P loop lengths, the relative curve for Type-II regulation is always higher than the ones for two other regulations, and the change trend of the former is opposite to that of the latter two.

In order to verify the above qualitative results, we further analyze the dependence of the mean mRNA level on both E-P loop length and feedback strength, referring to Figure 5D. Interestingly, we observe that for any feedback strength, the mean mRNA level can reach a minimum at about $2000\,bp$ E-P loop length. Particularly in the cases of Type-II and mixed regulations, the larger feedback strength leads to a higher mean mRNA level, but the E-P loop length needs to be limited in the interval from 200 to $4000\,bp$.

It is worth pointing out that the above qualitatively different relative change tendencies may in turn help us infer the way and the qualitative effect of eRNA regulation.

3.4. Effects of the E-P loop length on the mRNA noise

Figure 6 shows how the E-P loop length affects the mRNA noise. Specifically, Figure 6A demonstrates dependence of the mRNA noise on the E-P loop length in the four cases of eRNA regulation as specified above; Figure 6B, an enlarged diagram for the red dotted line shown in Figure 6A, demonstrates the dependence of the mRNA noise on the E-P loop length in the case of Type-I regulation; Figure 6C demonstrates the dependence of the ratio of the mRNA noise (defined as the ratio of the mRNA noise in the case of eRNA regulation over the mRNA noise in the case of non-regulation) on the E-P loop length in the three cases of eRNA regulation; and Figure 6D demonstrates the dependence of mRNA noise on both E-P loop length and feedback strength.

From Figure 6A, we first observe that the curve in the case of non-regulation is above the curve in the case of Type-I regulation, which is above the curve in the case of mixed regulation, which is above the curve in the case of Type-II regulation. This implies that each of three eRNA regulation ways can reduce the mRNA noise in contrast to non-regulation. In addition, whatever the E-P loop length, the mRNA noise in the case of Type-II regulation is the lowest, the mRNA noise in the case of Type-I regulation is the highest, and the mRNA noise in the case of mixed regulation is moderate. Then, we observe that in all the three cases of eRNA regulation, the mRNA noise intensity first increases and then decreases with the increase of the E-P loop length, implying that the mRNA noise has the highest level even for Type-I regulation (referring to Figure 6B).

From Figure 6C, we observe that the noise-ratio curves for Type-I regulation (red) and mixed regulation (purple) have the same change trend, that is, they first decrease and then increase until they become flat. The noise ratios reach the minimum around the $2000\,bp$ E-P loop length. If the E-P loop is long enough (e.g., more than $6000\,bp$), the noise level in the case of Type-I regulation is very close to that in the case of non-regulation, and the former is much smaller than the latter otherwise. In addition, the curve for Type-II regulation is always below the ones for the other two regulation ways, and the change trend of the former is opposite to that of the latter two. Furthermore, we analyze the simultaneous dependence of mRNA noise intensity on E-P loop length and feedback strength, referring to Figure 6D. From this diagram, we observe that for any feedback strength, the mRNA noise intensity

reaches the maximum around the 2000 *bp* E-P loop length. Particularly in the cases of Type-I and mixed regulations, the greater the feedback strength is, the smaller is the mRNA noise, and this relationship holds for the E-P loop length in the interval from 200 to 4000 *bp*.

The combination of Figures 5 and 6 implies that the Type-II regulation is advantageous over Type-I or mixed regulation in controlling gene expression and tuning gene expression noise, and this advantage is independent of the E-P loop length.

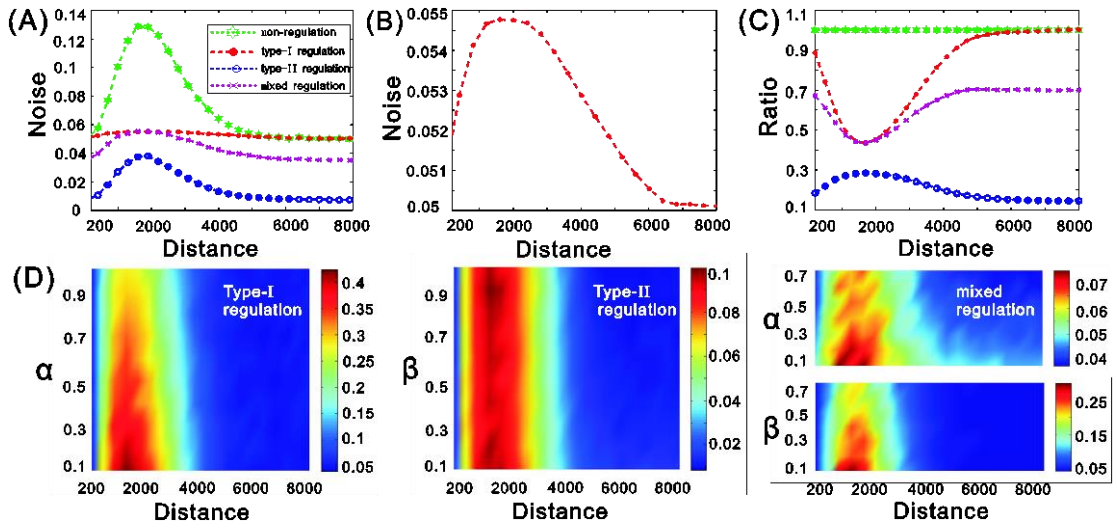

Figure 6. Influence of eRNA on the mRNA noise. (A) Dependence of the mRNA noise on the E-P loop length. (B) An enlarged diagram of the red dotted line in (A). (C) The ratio between the mRNA noise intensities in three cases of eRNA regulation. (D) Dependence of the mRNA noise on both E-P loop length and feedback strength. Parameter values are the same as those in Figure 5.

3.5. Effects of feedback strength and eRNA synthesis rate on mRNA noise

Here, we investigate the effects of feedback strength parameters (α and β) and eRNA production rate (μ_2) on the mRNA noise, for a fixed E-P loop length. Numerical results are demonstrated in Figure 7.

From Figure 7A and B, we observe that the mRNA noise is a monotonously decreasing function of feedback strength α or β, where red dotted lines correspond to Type-I regulation, blue circled lines to Type-II regulation, and purple cross lines to mixed regulation. Figure 7A shows that with the increase of α, the mRNA noise for Type-I regulation is a bit greater than that for mixed regulation. Similarly, Figure 7B shows that with the increase of β, the mRNA noise for Type-II regulation is

apparently greater than that for mixed regulation. In addition, the mRNA noise in Figure 7A is a concave function of α, whereas that in Figure 7B has an inflecting point nearly at $\beta = 0.45$: the mRNA noise is a concave function if $\beta < 0.45$ but a convex function if $\beta > 0.45$.

Figure 7C demonstrates how the mRNA noise intensity changes with the eRNA synthesis rate μ_2. We observe that in all the three cases of eRNA regulation, with the increasing of synthesis rate μ_2, the mRNA noise first decreases rapidly and then maintains at a stable level after the synthesis rate is beyond a certain value. In the three cases of eRNA regulation, although the change trend of the mRNA noise is fundamentally similar, the mRNA noise level displays an apparent difference. Specifically, if μ_2 is kept the same, the line for Type-I regulation is above that for mixed regulation, which is above that for Type-II regulation. This indicates that if the eRNA generation rate is set as unit, the mRNA noise for Type-II regulation is the lowest, and the mRNA noise for Type-I regulation is higher than that for mixed regulation. Therefore, from the viewpoint of reducing the mRNA noise, the best way of eRNA regulation is the Type-II regulation, which is in accordance with the conclusion for the mean mRNA.

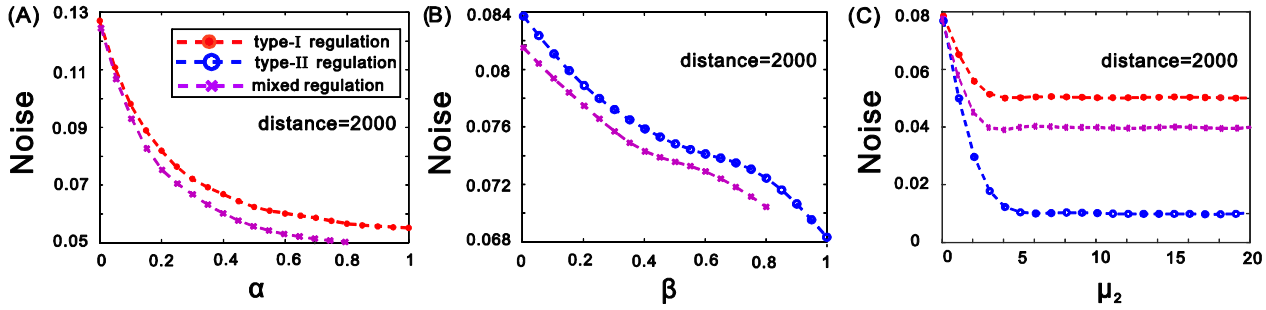

Figure 7. Dependence of the mRNA noise on feedback strength or the eRNA synthesis rate in three cases of eRNA regulation, where red dotted lines correspond to Type-I regulation, blue circled lines to Type-II regulation, and purple cross lines to mixed regulation. (A) Effect of α on mRNA noise, where α changes from 0 to 1 for Type-I regulation, and from 0 to 0.8 for mixed regulation with the fixed $\beta = 0.2$. (B) Effect of β on mRNA noise, where β changes from 0 to 1 for Type-II regulation, and from 0 to 0.8 for mixed regulation with the fixed $\alpha = 0.2$. (C) Effect of μ_2 on mRNA noise, where parameter μ_2 changes from 1 to 20, $\alpha = 0.1$ and $\beta = 0.9$ for mixed regulation. In (A), (B) and (C), parameter values are set as: $k_1 = 1, \mu_1 = 20, \delta_1 = 1, \delta_2 = 5, d = 2000bp$.

4. Conclusions

More and more experimental evidences indicate that non-coding RNAs play important role in controlling gene expression. As a kind of mysterious non-coding RNA and a new regulation factor, eRNA may affect gene expression and its anomalous expression may lead to many diseases [58–61,34]. Exploratory interrogation should be conducted to unravel the association of functional eRNAs with specific chromatin regions and other functional RNAs [62], as well as to understand possible functional motifs, secondary structures and chemical modifications [63,64]. On the other hand, mathematical models are a powerful tool for exploring possible functions of eRNAs.

In this work, based on biologically experimental evidence, we have proposed a mechanistic model of stochastic gene expression with eRNA regulation. In this toy model, the eRNA regulatory role is primarily reflected in one of three ways: (1) Type-I regulation by which eRNA accelerates the formation of the E-P loop, (2) Type-II regulation by which eRNA promotes the production rate of its target gene product, (3) mixed regulation which is a combination of the former two regulation ways. We have also provided a theoretical analysis framework to investigate the eRNA's contribution to the target gene expression, and derived analytical formulae for the mRNA's distribution, mean and noise. Interestingly, we have found that with the increase of the E-P loop length, the mRNA distributions can transition to bimodality from unimodality and then be back to unimodality, independent of eRNA regulation ways. We have also shown that in contrast to Type-I and mixed regulations, the Type-II regulation can result in the higher mean mRNA level and the lower mRNA noise, independent of the E-P loop length. These results are independent of the choice of the model's parameter values and are hence qualitative, although we only demonstrated relevant numerical results for particular sets of parameter values. Our results indicate that the E-P loop can induce phenotypic switching, a mechanism different from the traditional one (i.e., feedback can induce phenotypic switching). Both would be exploited by live cells surviving in complex environments.

In our model, eRNA is taken as an enhancer's product, which inter-regulates the expression of its target gene. The made assumption may not be exactly consistent with biological reality of eRNA regulation under some specific environments. In other situations, it would be needed to construct other more biologically reasonable models to investigate the role of eRNA in regulating its target genes, e.g., the constructed models should consider dynamic regulation where eRNA is taken as an input signal of the model systems [65]. For this, one may assume that eRNA is a special spatiotemporal regulation factor that wanders in cellular space, and its concentration is a time-dependent function, eRNA(t). Based on this assumption, one can also establish theoretical models as done in this paper. This will be the next task in our future research.

It should be pointed out that in the introduction, we have mentioned three hypotheses on the "functional theory": eRNA facilitates the formation of the E-P loop, eRNA remodels chromatin, and eRNA improves the extension of RNAPII. In this paper, we assumed that the most possible biological function of eRNA is to enhance the mRNA production rate. Although this rate integrates the effects of chromatin remodeling and promoting RNAPII elongation, but we did not specify molecular mechanisms behind these effects. Therefore, the future study needs to construct a more realistic model to investigate whether chromatin remodeling or the extension of RNAIIP can increase the mRNA transcriptional rate, and how this rate affects gene expression and further cell-to-cell variability across a genetically identical population of cells.

Acknowledgments

This paper was supported by grant 11775314 and 11931019 (TS), 11801482(HL), 11701591(TQ) from Natural Science Foundation of P. R. China; grant 202007030004 from Key-Area Research and Development Program of Guangzhou.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. T. Maniatis, S. Goodbourn, J. A. Fischer, Regulation of inducible and tissue -specific gene expression, *Science*, **236** (1987), 1237–1245. doi: 10.1126/science.3296191.
2. E. M. Blackwood, J. T. Kadonaga, Going the Distance: A Current View of Enhancer Action, *Science*, **281** (1998), 60–63. doi: 10.1126/science.281.5373.60.
3. L. A. Pennacchio, W. Bickmore, A. Dean, M. A. Nobrega, G. Bejerano, Enhancers: five essential questions, *Nat. Rev. Genet.*, **14** (2013), 288–295. doi: 10.1126/science.281.5373.60
4. G. A. Maston, S. K. Evans, M. R. Green, Transcriptional Regulatory Elements in the Human Genome, *Annu. Rev. Genom. Hum. G.*, **7** (2006), 29–59. doi: 10.1146/annurev.genom.7.080505.115623.
5. O. I. Kulaeva, E. V. Nizovtseva, Y. S. Polikanov, S. V. Ulianov, V. M. Studitsky, Distant activation of transcription: mechanisms of enhancer action, *Mol. Cell Biol.*, **32** (2012), 4892–4897. doi: 10.1128/MCB.01127-12.
6. C. Buecker, J. Wysocka, Enhancers as information integration hubs in development: lessons from genomics, *Trends Genet.*, **28** (2012), 276–284. doi: 10.1016/j.tig.2012.02.008.
7. W. Xie, B. Ren, Enhancing Pluripotency and Lineage Specification, *Science,* **341** (2013), 245-247. doi: 10.1126/science.1236254.
8. K. Lee, C. C. Hsiung, P. Huang, A. Raj, G. A. Blobel, Dynamic enhancer–gene body contacts during transcription elongation, *Genes. Dev.*, **29** (2015), 1992–1997. doi: 10.1101/gad.255265.114.
9. J. G. Azofeifa, M. A. Allen, J. R. Hendrix, T. Read, J. D. Rubin, R. D. Dowell, Enhancer RNA profiling predicts transcription factor activity, *Genome Res.*, **28** (2018), 334–344. doi: 10.1101/gr.225755.117.
10. E. Zlotorynski, Gene expression: Developmental enhancers in action, *Nat. Rev. Genet.*, **19** (2018), 187–187. doi: 10.1038/nrg.2018.13.
11. A. Smallwood, B. Ren, Genome organization and long-range regulation of gene expression by enhancers, *Curr. Opin. Cell Biol.*, **25** (2013), 387–394. doi: 10.1016/j.ceb.2013.02.005.
12. A. Maruyama, J. Mimura, K. Itoh, Non-coding RNA derived from the region adjacent to the human HO-1 E2 enhancer selectively regulates HO-1 gene induction by modulating Pol II binding, *Nucleic Acids Res.*, **42** (2014), 13599–13614. doi: 10.1093/nar/gku1169.
13. C. L. Yin, C. Benner, R. Mansson, S. Heinz, K. Miyazaki, M. Miyazaki，et al., Global changes in nuclear positioning of genes and intra- and inter-domain genomic interactions that orchestrate B cell fate, *Nat. Immunol.*, **13** (2012), 1196–1204. doi: 10.1038/ni.2432.

14. A. Sanyal, B. R. Lajoie, G. Jain, J. Dekker, The long-range interaction landscape of gene promoters, *Nature* **489** (2012),109–113. doi: 10.1038/nature11279.
15. G. Natoli, J. Andrau, Noncoding transcription at enhancers: general principles and functional models, *Annu. Rev. Genet.*, **46** (2012), 1–19. doi: 10.1146/annurev-genet-110711-155459.
16. M. D. Young, T. A. Willson, M. J. Wakefield, T. Evelyn, D. J. Hilton, M. E. Blewitt, et al., Chip-seq analysis reveals distinct h3k27me3 profiles that correlate with transcriptional activity, *Nucleic Acids Res.*, **39** (2011), 7415–7427. doi: 10.1093/nar/gkr416.
17. F. Koch, R. Fenouil, M. Gut, P. Cauchy, T. K. Albert, J. Za Ca Rias-Cabeza, et al., Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters, *Nat. Struct. Mol. Bio.*, **18 (**2011), 956–963. doi: 10.1038/nsmb.2085.
18. M. S. Kowalczyk, J. R. Hughes, D. Lynch, M. D. Garrick, D. R. Higgs, Intragenic enhancers act as alternative promoters, *Mol. Cell*, **45** (2012), 447–458. doi: 10.1016/j.molcel.2011.12.021.
19. D. Wang, I. Garcia-Bassets, C. Benner, W. Li, S. Xue, Y. Zhou, et al., Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA, *Nature*, **474** (2011), 390–394. doi: 10.1038/nature10006.
20. J. Lacava, J. Houseley, C. Saveanu, E. Petfalski, E. Thompson, A. Jacquier, et al., RNA degradation by the exosome is promoted by a nuclear polyadenylation complex, *Cell*, **121** (2005), 713–724. doi: 10.1016/j.cell.2005.04.029.
21. F. Wyers, M. Rougemaille, G. Badis, J. C. Rousselle, A. Jacquier, Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase, *Cell*, **121** (2005), 725–737. doi: 10.1016/j.cell.2005.04.030.
22. I. Miguel-Escalada, R. Andersson, E. Al, C. Gebhard, I. Hoof, J. Bornholdt, An atlas of active enhancers across human cell types and tissues, *Nature*, **507** (2014), 455–461. doi: 10.1038/nature12787.
23. R. A. Flynn, A. E. Almada, J. R. Zamudio, P. A. Sharp, Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome, *Proc. Natl. Acad. Sci. USA.*, **108** (2011), 10460–10465. doi: 10.1073/pnas.1106630108.
24. N. Hah, S. Murakami, A. Nagari, C. G.Danko, W. L. Kraus, Enhancer transcripts mark active estrogen receptor binding sites, *Genome Res.*, **23** (2013), 1210–1223. doi: 10.1101/gr.152306.112.
25. A. Sandelin, P. Carninci, B. Lenhard, J. Ponjavic, Y. Hayashizaki, D. A. Hume, Mammalian RNA polymerase II core promoters: insights from genome-wide studies, *Nat. Rev. Genet.*, **8** (2007), 424–436. doi: 10.1038/nrg2026.
26. Struhl, Kevin, Transcriptional noise and the fidelity of initiation by RNA polymerase II, *Nat. Struct. Mol. Bio.*, **14** (2007), 103–105. doi: 10.1038/nsmb0207-103.
27. W. Li, D. Notani, Q. Ma, B. Tanasa, E. Nunez, A. Y. Chen, et al., Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation, *Nature*, **498** (2013), 516–520. doi: 10.1038/nature12210.
28. M. T. Y. Lam, C. Han, H. P. Lesch, D. Gosselin, S. Heinz, Y. Tanaka-Oishi, et al., Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription, *Nature*, **498** (2013)**,** 511–515. doi: 10.1038/nature12209.
29. C. L. Hsieh, T. Fei, Y. Chen, T. Li, Y. Gao, X. Wang, et al., Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation, *Proc. Natl. Acad. Sci. USA.*, **111** (2014),7319–7324. doi: 10.1073/pnas.1324151111.

30. C. Melo, J. Drost, P. Wijchers, H. Van De Werken, E. De Wit, J. F. Vrielink, et al., eRNAs are required for p53-dependent enhancer activity and gene transcription, *Mol. Cell*, **49** (2013), 524–535. doi: 10.1016/j.molcel.2012.11.021.
31. K. Mousavi, H. Zare, S. Dell'orso, L. Grøntved, G. Gutierrez-Cruz, A. Derfoul, et al., eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci, *Mol. Cell*, **51** (2013), 606–617. doi: 10.1016/j.molcel.2013.07.022.
32. N. E. Ilott, J. A. Heward, B. Roux, E. Tsitsiou, P. S. Fenwick, L. Lenzi, et al., Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes, *Nat. Commun.*, **5** (2014), 3979–3979. doi: 10.1038/ncomms4979.
33. K. Schaukowitch, J. Y. Joo, X. Liu, J. K. Watts, C. Martinez, T. K. Kim, Enhancer RNA facilitates NELF release from immediate early genes, *Mol. Cell*, **56** (2014), 29–42. doi: 10.1016/j.molcel.2014.08.023.
34. T. K. Kim, R. Shiekhattar, Architectural and functional commonalities between enhancers and promoters, *Cell*, **162** (2015), 948–959. doi: 10.1016/j.cell.2015.08.008.
35. E. Savitskaya, L. Melnikova, M. Kostuchenko, E. Kravchenko, E. Pomerantseva, T. Boikova, et al., Study of long-distance functional interactions between Su(Hw) insulators that can regulate enhancer-promoter communication in Drosophila melanogaster, *Mol. Cell Biol.*, **26** (2006), 754–761. doi: 10.1128/MCB.26.3.754-761.2006.
36. D. Chetverina, M. Fujioka, M. Erokhin, P. Georgiev, J. B. Jaynes, P. Schedl, Boundaries of loop domains (insulators): Determinants of chromosome form and function in multicellular eukaryotes, *Bioessays.*, **39** (2017), 10.1002. doi: 10.1002/bies.201600233.
37. J. Mozziconacci, M. Merle, A. Lesne, The 3D Genome Shapes the Regulatory Code of Developmental Genes, *J. Mol. Biol.*, **432** (2020),712–723. doi: 10.1016/j.jmb.2019.10.017.
38. C.P. Fulco, J. Nasser, T. R. Jones, G. Munson, D.T. Bergman, V. Subramanian, et al., Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations, *Nat Genet*., **51** (2019):1664–1669. doi: 10.1038/s41588-019-0538-0.
39. T. Q. Liu, J. J. Zhang, T. S. Zhou, Effect of Interaction between Chromatin Loops on Cell-to-Cell Variability in Gene Expression, *PLoS Comput. Biol.*, **12** (2016), e1004917. doi: 10.1371/journal.pcbi.1004917.
40. J. J. Zhang, T. S. Zhou, Promoter architecture-mediated transcriptional dynamics, *Biophys. J.*, **106** (2014), 479–488. doi: 10.1016/j.bpj.2013.12.011.
41. J. Peccoud, B. Ycart, Markovian Modeling of Gene-Product Synthesis, *Theor. Popul. Biol.*, **48** (1995), 222–234. doi: 10.1006/tpbi.1995.1027.
42. A. F. Ramos, G. Innocentini, F. M. Forger, J. Hornos, Symmetry in biology: from genetic code to stochastic gene regulation, *Iet Syst. Biol.*, **4** (2010), 311–329. doi: 10.1049/iet-syb.2010.0058.
43. G. Giovanini, A. U. Sabino, L. R. C. Barros, A. F. Ramos, A comparative analysis of noise properties of stochastic binary models for a self-repressing and for an externally regulating gene, *Math. Biosci. Eng.*, **17** (2020), 5477–5503. doi: 10.3934/mbe.2020295.
44. N. Kumar, T. Platini, R. V. Kulkarni, Exact distributions for stochastic gene expression models with bursting and feedback, *Phys. Rev. Lett.*, **113** (2014), 268105–268105. doi: 10.1103/PhysRevLett.113.268105.
45. R. Grima, D. R. Schmidt, T. J. Newman, Steady-state fluctuations of a genetic feedback loop: An exact solution. *J. Chem. Phys.*, **137** (2012), 051907. doi: 10.1063/1.4736721.

46. G. Innocentini, A. F. Ramos, J. Hornos, Comment on "Steady-state fluctuations of a genetic feedback loop: an exact solution" [J. Chem. Phys. {\bf 137, 035104 (2012).], *J. Chem. Phys.*, **142** (2015), 027101. doi: 10.1063/1.4905217.
47. A. Sanchez, H. G. Garcia, D.Jones, R. Phillips, J. Kondev, Effect of promoter architecture on the cell-to-cell variability in gene expression, *Plos Comput. Biol.*, **7**(2011), e1001100. doi: 10.1371/journal.pcbi.1001100.
48. L. Bintu, N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev, et al., Transcriptional regulation by the numbers: applications, *Curr. Opin. Genet. De.*, **15** (2005), 125–135. doi: 10.1016/j.gde.2005.02.006.
49. H. Yamakawa, *In Helical Wormlike Chains in Polymer Solutions*, Springer Press, New York, 1997. doi: 10.1007/978-3-642-60817-9.
50. F. De Santa, I. Barozzi, F. Mietton, S. Ghisletti, S. Polletti, B. K. Tusi, et al., A large fraction of extragenic RNA pol II transcription sites overlap enhancers, *PLoS Biol.*, **8** (2010), e1000384. doi: 10.1371/journal.pbio.1000384.
51. T. K. Kim, M. Hemberg, J. M. Gray, A. M. D. M. Costa, J. Bear, D. Wu, et al., Widespread transcription at neuronal activity-regulated enhancers, *Nature*, **465** (2010), 182–187. doi: 10.1038/nature09033.
52. L. J. Slater, *Confluent Hypergeometric Functions*, Cambridge University Press, Cambridge, 1960.
53. D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81** (1977), 2340–2361. doi: 10.1021/j100540a008.
54. D. Fraser, M. Kaern, A chance at survival: gene expression noise and phenotypic diversification strategies, *Mol. Microbiol.*, **71** (2009), 1333–1340. doi: 10.1111/j.1365-2958.2009.06605.x.
55. M. Acar, J. T. Mettetal, A. Van Oudenaarden, Stochastic switching as a survival strategy in fluctuating environments, *Nat. Genet.*, **40** (2008), 471–475. doi: 10.1038/ng.110.
56. S. Iyer-Biswas, F. Hayot, C. Jayaprakash, Stochasticity of gene products from transcriptional pulsing, *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, **79** (2009), 031911. doi: 10.1103/PhysRevE.79.031911.
57. R. Hermsen, D. W. Erickson, T. Hwa, Speed, sensitivity, and bistability in auto-activating signaling circuits, *Plos Comput. Biol.*, **7** (2011), e1002265. doi: 10.1371/journal.pcbi.1002265.
58. E. Lenters-Westra, T. Røraas, R. K. Schindhelm, R. J. Slingerland, S. Sandberg, Biological variation of hemoglobin A_{1c}: consequences for diagnosing diabetes mellitus, *Clin. Chem.*, **60** (2014), 1570–1572. doi: 10.1373/clinchem.2014.227983.
59. L. Pnueli, S. Rudnizky, Y. Yosefzon, P. Melamed, RNA transcribed from a distal enhancer is required for activating the chromatin at the promoter of the gonadotropin α-subunit gene, *Proc. Natl. Acad. Sci. USA.*, **112** (2015), 4369–4374. doi: 10.1073/pnas.1414841112.
60. C. Benner, T. Isoda, C. Murre, New roles for DNA cytosine modification, eRNA, anchors, and superanchors in developing B cell progenitors, *Proc. Natl. Acad. Sci. USA.*, **112** (2015), 12776–12781. doi: 10.1073/pnas.1512995112.
61. D. Holoch, D. Moazed, RNA-mediated epigenetic regulation of gene expression, *Nat. Rev. Genet.*, **16** (2015), 71–84. doi: 10.1038/nrg3863.
62. J. M. Engreitz, K. Sirokman, P. McDonel, A. A., Shishkin, C. Surka, P. Russell, et al., RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites, *Cell*, **159** (2014),188–199. doi: 10.1016/j.cell.2014.08.018.

63. J. Quinn, H. Chang, Unique features of long non-coding RNA biogenesis and function, *Nat. Rev. Genet.*, **17** (2016), 47–62. doi: 10.1038/nrg.2015.10.
64. F. Aguilo, S. Li, N. Balasubramaniyan, A. Sancho, S. Benko, F. Zhang, et al., Deposition of 5-methylcytosine on enhancer RNAs enables the coactivator function of PGC-1α, *Cell Rep.*, **14** (2016), 479–492. doi: 10.1016/j.celrep.2015.12.043.
65. W. Li, D. Notani, M. G. Rosenfeld, Enhancers as non-coding RNA transcription units: recent insights and future perspectives, *Nat. Rev. Genet.*, **17** (2016), 207–223. doi: 10.1038/nrg.2016.4.

Appendix A: Derivation of analytical results

Let $Q_i(n)=\sum_m P_i(m,n)\quad (i=I,A)$ be the factorial marginal probability density function of eRNA. It follows from Eq 3 in the main text that

$$\begin{aligned} &k_1 Q_A(n)-(k_0+\alpha n)Q_I(n)+\delta_2(E_n-S)\left[nQ_I(n)\right]=0\\ &-k_1 Q_A(n)+(k_0+\alpha n)Q_I(n)+\delta_2(E_n-S)\left[nQ_A(n)\right]+\mu_2\left(E_n^{-1}-S\right)\left[Q_A(n)\right]=0 \end{aligned} \tag{A1}$$

This is a solvable model where mRNA does not impact eRNA. To solve this equation group, we introduce probability-generating functions $G_i(z)=\sum_n (z+1)^n Q_i(n), i=I,A,$ and $G(z)=G_I(z)+G_A(z)$. Then

$$\begin{aligned} &\tilde{k}_1 G_A(z)-\tilde{k}_0 G_I(z)-\tilde{\alpha}(z+1)\partial G_I(z)-z\partial G_I(z)=0\\ &-\tilde{k}_1 G_A(z)+\tilde{k}_0 G_I(z)+\tilde{\alpha}(z+1)\partial G_I(z)-z\partial G_A(z)+\tilde{\mu}_2 z G_A(z)=0 \end{aligned} \tag{A2}$$

where the parameters are normalized by δ_2. By adding the above two equations, we get $G_A(z)=\frac{1}{\tilde{\mu}_2}\partial G(z)$ and $G_I(z)=G(z)-G_A(z)=G(z)-\frac{1}{\tilde{\mu}_2}\partial G(z)$. From the first equation of (A2), we have

$$\left[(\tilde{\alpha}+1)z+\tilde{\alpha}\right]\frac{d^2G(z)}{dz^2}-\left[\tilde{\mu}_2(\tilde{\alpha}+1)z+\tilde{\alpha}\tilde{\mu}_2-\tilde{k}_0-\tilde{k}_1\right]\frac{dG(z)}{dz}-\tilde{\mu}_2\tilde{k}_0 G(z)=0\,. \tag{A3}$$

That is,

$$(\tilde{\alpha}+1)\left(z+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}\right)\frac{d^2G(z)}{dz^2}-\tilde{\mu}_2(\tilde{\alpha}+1)\left[z+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}-\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\mu}_2(\tilde{\alpha}+1)}\right]\frac{dG(z)}{dz}-\tilde{\mu}_2\tilde{k}_0 G(z)=0\,. \tag{A4}$$

Setting $x=z+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}$, we have

$$x\frac{d^2G(x)}{dx^2}-\tilde{\mu}_2\left[x-\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\mu}_2(\tilde{\alpha}+1)}\right]\frac{dG(x)}{dx}-\frac{\tilde{\mu}_2\tilde{k}_0}{\tilde{\alpha}+1}G(x)=0\,. \tag{A5}$$

Making Taylor expansion: $G(x)=\sum_n a_n x^n$, yields the following iterative equation

$$n(n+1)a_{n+1}-\tilde{\mu}_2 n a_n+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1}(n+1)a_{n+1}-\frac{\tilde{\mu}_2\tilde{k}_0}{\tilde{\alpha}+1}a_n=0\,, \tag{A6}$$

which gives

$$a_n = \frac{\tilde{\mu}_2^n}{n!}\frac{\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1}\right)_n}{\left(\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1}\right)_n}a_0\,,\quad n=0,1,2,\cdots \tag{A7}$$

Thus, we obtain the following expression of function $G(x)$

$$G(x) = a_0\,{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\tilde{\mu}_2 x\right). \tag{A8}$$

Being back to the original variable yields

$$G(z) = G(x) = a_0\,{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\tilde{\mu}_2\left(z+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}\right)\right) \tag{A9}$$

Note that $G(0)=1$, which can give

$$a_0 = \left[{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)\right]^{-1} \tag{A10}$$

Using the relationship between probability distribution and probability-generating function, we thus obtain the analytical expression for the marginal probability density function of eRNA

$$Q(n) = \frac{1}{n!}\frac{d^n}{dz^n}G(z)\Big|_{z=0} = \frac{\tilde{\mu}_2^n}{n!}\frac{\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1}\right)_n}{\left(\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1}\right)_n}\frac{{}_1F_1\left(n+\frac{\tilde{k}_0}{\tilde{\alpha}+1},n+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)}{{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)} \tag{A11}$$

Furthermore, the mean of eRNA is given by

$$\langle n\rangle = \frac{dG(z)}{dz}\Big|_{z=1} = \frac{\tilde{\mu}_2\tilde{k}_0}{\tilde{k}_0+\tilde{k}_1}\frac{{}_1F_1\left(1+\frac{\tilde{k}_0}{\tilde{\alpha}+1},1+\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\tilde{\mu}_2\left(1+\frac{\tilde{\alpha}}{\tilde{\alpha}+1}\right)\right)}{{}_1F_1\left(\frac{\tilde{k}_0}{\tilde{\alpha}+1},\frac{\tilde{k}_0+\tilde{k}_1}{\tilde{\alpha}+1};\frac{\tilde{\mu}_2\tilde{\alpha}}{\tilde{\alpha}+1}\right)} \tag{A12}$$

Next, let $R_i(m)=\sum_n P_i(m,n)$ be the marginal probability density functions of mRNA. It follows from Eq 3 in the main text that

$$\begin{aligned}
&\tilde{k}_1 R_A(m) - \tilde{k}_0 R_I(m) - \alpha\sum_{n=0}^{\infty} nP_I(m,n) + (E_m - S)\left[mR_I(m)\right] = 0\\
&-\tilde{k}_1 R_A(m) + \tilde{k}_0 R_I(m) + \alpha\sum_{n=0}^{\infty} nP_I(m,n) + (E_m - S)\left[mR_A(m)\right]\\
&\quad + \left(\tilde{\mu}_1 + \tilde{\beta}\langle n\rangle\right)\left(E_m^{-1} - S\right)\left[R_A(m)\right] = 0
\end{aligned} \tag{A13}$$

Now, we make the approximation: replacing $P_I(m,n)$ with $R_I(m)Q_I(n)$. In addition, we make the approximation: replacing $\sum_{n=0}^{\infty} nQ_I(n)$ with $\sum_{n=0}^{\infty} nQ(n).$ The effectiveness of these approximations has been verified by numerical simulation (seeing the main text). Then, we have the approximate equations

$$\begin{aligned}&\tilde{k}_1 R_A(m)-\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle\right)R_I(m)+\left(E_m-S\right)\left[mR_I(m)\right]\approx 0\\&-\tilde{k}_1 R_A(m)+\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle\right)R_I(m)+\left(E_m-S\right)\left[mR_A(m)\right]+\left(\tilde{\mu}_1+\tilde{\beta}\langle n\rangle\right)\left(E_m^{-1}-S\right)\left[R_A(m)\right]\approx 0\end{aligned} \tag{A14}$$

where $\langle n\rangle$ is given by Eq A12. Since $\langle n\rangle$ is unrelated to mRNA, it can be taken as a parameter in Eq A14. The other parameters in Eq A14 have been normalized by δ_1. Thus, Eq A14 is also a solvable model. In fact, it is a common ON-OFF model where the transition rate from ON to OFF state is $\tilde{k}_1$, the transition rate from OFF to ON state is $\tilde{k}_0+\tilde{\alpha}\langle n\rangle$ and mRNA transcriptional rate is $\tilde{\mu}_1+\tilde{\beta}\langle n\rangle$. By solving this model, we obtain the following analytical expression of stationary mRNA probability distribution

$$R(m)=\frac{\left(\tilde{\mu}_1+\tilde{\beta}\langle n\rangle\right)^m}{m!}\frac{\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle\right)_m}{\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle+\tilde{k}_1\right)_m}\,{}_1F_1\left(\tilde{k}_0+\tilde{\alpha}\langle n\rangle+m,\tilde{k}_0+\tilde{\alpha}\langle n\rangle+\tilde{k}_1+m;-\left(\tilde{\mu}_1+\tilde{\beta}\langle n\rangle\right)\right) \tag{A15}$$

where $(c)_n$ is the Pochhammer symbol, defined as $(c)_n=\Gamma(c+n)/\Gamma(n)$, and ${}_1F_1(a,b;z)$ is a confluent hypergeometric function, defined as ${}_1F_1(a,b;z)=\sum_{n=0}^{\infty}\frac{(a)_n}{(b)_n}\frac{z^n}{n!}$.

Appendix B: Supplemental figure

In the main text, we have investigated the effect of eRNA on the mRNA expression in the case that eRNA regulates the transition rate from OFF to ON. However, it is possible that eRNA regulates the transition rate from ON to OFF. In this case, whether the qualitative effect of eRNA affects on the mRNA expression is that in the case that eRNA regulates the transition rate from OFF to ON. We numerically find that both are indeed the same. Here we only demonstrate an example, referring to Figure S1.

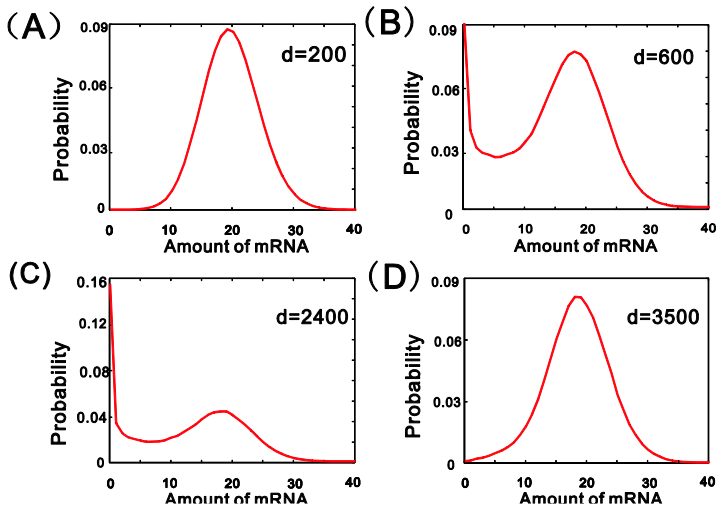

Figure S1. Influence of reducing k_1 on stationary mRNA distribution, where distance d is determined by $k_1 = k_{on}\exp\left(-\frac{u}{d} - v\log d + wd + z\right)$, and some parameter values are set as: $k_0 = 1, \mu_1 = 20, \mu_2 = 2, \delta_1 = 1, \delta_2 = 5, \alpha = 1.$